Received: 11/06/2025, Accepted: 23/07/2025

Research Article

Smart Detection of Credit Card Fraud Using Machine Learning

Amit Bhola, Vinayak Pandey, Ayush Kumar, Prashant Tomar
Department of Computer Science and Engineering, Sharda University, Greater Noida, India
amitbhola20@gmail.com, pandeyvinayak076@gmail.com, pennywise024680@gmail.com, prashantbahadur2003@gmail.com

Abstract

The increase in digital financial transactions has dramatically increased credit card fraud, which now poses significant threats to consumers as well as financial institutions. This study provides a comparative study of a variety of machine learning models for credit card fraud detection using a real life, imbalanced dataset. We compared four machine learning models—Logistic Regression (LR), Random Forest (RF), XGBoost, and Artificial Neural Network (ANN)—in terms of their ability to detect fraudulent transactions accurately. We used the Synthetic Minority Over-sampling Technique (SMOTE) for preprocessing and balancing the datasets, but tested the machine learning models on the original imbalanced set to adequately represent real-world performance. The experimental results of this study indicate that the ANN and XGBoost models were the most accurate and had the highest recall and F1-score, with the ANN performing the best in most corrective metrics. The feature importance plot shows that there are some PCA variables, such as V14 and V17, that are relevant to fraudulent activity. The research provides evidence of the capabilities of ensemble and deep learning models in performing fraud detection tasks, especially after careful preprocessing and addressing problems related to the dataset imbalance. This study also provides a computationally sound and useful methodology that could be utilized in developing intelligent fraud detection systems in the financial sector.

Keywords: Machine learning, XGBoost, Artificial Neural Network, Imbalanced classification, SMOTE

1 Introduction

In today's fast paced digital economy, and credit cards have emerged as a key and core medium for financial transactions around the globe. From e-commerce platforms to contactless in-store payments, credit cards offer an easy, flexible, and convenient way of transacting. However, this increased dependency exposes users and financial institutions to an increased number of opportunistic fraudulent activities. Credit card fraud generates significant monetary losses on behalf of all parties involved and affects consumer trust in the financial ecosystem.

A report from Nilson [1] estimated that global losses from credit card fraud was greater than $32 billion in 2022 and is expected to be more than $40 billion by 2027. The report illustrates the critical need to automate fraud detection systems and develop more robust and smart solutions for fraud detection. The problem is that fraudulent transactions often behave like genuine transactions, making it nearly impossible to detect fraud using traditional rules-based approaches.

For a long time, rule-based systems have provided the foundation for fraud detection infrastructures. Rule-based systems consist of rules that are made manually to flag transactions based on static conditions tied to the parameters of the payment transaction (e.g., uncharacteristic location, high dollar amount, irregular purchasing frequency). While rule-based systems are good at catching known patterns of fraud, they lack the flexibility to adaptive to changing fraudster behaviors over time. Additionally, rule-based systems have high false positive rates trigger excessive alert cycles and poor customer experience. As fraudsters continuously create unambiguous ways to bypass checks & balances, rule-based systems are ineffectively making obsolete.

*Corresponding author: Department of Computer Science and Engineering, Sharda University, Greater Noida, India (amitbhola20@gmail.com)

To address these constraints, it means already considering machine learning (ML) techniques is a formidable alternative to combat credit card fraud. Machine learning allows systems to learn from history and identify complex patterns that would be difficult to capture manually. By monitoring transaction data and automatically observing correlations and anomalies, ML models improve both accuracy and efficiency in detecting fraud [2]. One of the distinguishing features of machine learning is its flexibility. Fraudulent behavior changes constantly, and the models can be retrained as those changes occur. Furthermore, ML-based systems can ingest large quantities of transactional data as it flows in real time, making them consistent with transacting in volatile financial environments. Yet, applying ML to fraud detection brings with it a multitude of challenges that need to be carefully managed.

One of the most notable problems centers around the issue of class imbalance. In a typical credit card transaction dataset, not only fraudulent transactions are typically only a tiny fraction of transactions - often less than 0.2% of the total data. This class imbalance can bias models towards the majority class (legitimate transactions) and limit the ability of the model to detect the minority class (frauds). To combat class imbalance, many researchers use techniques such as the Synthetic Minority Over-sampling Technique (SMOTE), or alternatively undersampling or relying on a combination of both techniques [3]. Another factor to consider is which algorithms to use. The multitude of machine learning models available to researchers has different advantages and disadvantages that depend on the data you will be working with. For example, logistic regression offers a level of interpretability and simple models; however, it does not perform well if the pattern is highly nonlinear. One of the drawbacks of ensemble methods, for instance Random Forest and XGBoost, is that they are robust and highly accurate, but they are computationally expensive and will sacrifice some interpretability. Deep learning methods, like artificial neural networks (ANNs), are effective and truly shine for high-dimensional data. However, they tend to be slow to train and costly resource-wise [4].

Besides accuracy, other metrics such as precision, recall, F1-score, and the area under the Receiver Operating Characteristic (ROC-AUC) curve should also be evaluated in fraud detection systems. Recall is a particularly important metric in fraud detection because failing to detect a fraudulent transaction can have major consequences. While recall tends to be the priority, equally one must take extra precautions to have a low misclassification rate. If a system is too sensitive, many legitimate transactions may be flagged, leading consumers to become annoyed with transaction notifications and financial institutions becoming overwhelmed by paperwork.

This paper provides a detailed exploration of four machine learning algorithms for credit card fraud detection—Logistic Regression, Random Forest, XGBoost and Artificial Neural Networks. We utilize a publicly available dataset that reflects a real-world distribution of fraud, use of data preprocessing techniques (data cleaning), and use SMOTE to address the class imbalances in the data. We evaluate all four algorithms using a variety of performance metrics. As an added value to practitioners, we provide visual representations of the results with the utilization of tables and Receiver Operating Characteristic (ROC) curves that highlight the performance and limitations of the algorithms.

The purpose of this research is not simply to identify the best model, but also to develop a usable framework for implementing machine learning-based fraud detection systems for financial applications. We stress the balance between model performance, computation efficiency, and real-time applicability. Hopefully, our findings add to the growing body of work on intelligent fraud prevention and assist financial institutions to develop more secure and integrated methods of transactions.

2 Literature Review

Credit card fraud detection has hosted a long line of research for about twenty years now. Traditionally it was a rule-based system, but in recent years it has transitioned to data-based systems that are more driven by intelligence. More specifically, machine learning (ML) and deep learning, has received attention because it scales well with larger volume transaction data and recognizes complex or sophisticated fraud patterns that are not easy to define by hand. This section wraps up a critical survey of recent literature citing recent patterns and results that are pertinent to the current state of fraud detection using machine learning.

2.1 Ensemble Learning Models

Ensemble learning techniques such as Random Forest and Gradient Boosted Decision Trees (GBDTs) have been successfully applied in the field of fraud detection. Khatri et al. [5] performed a comparative study on real-world transaction data and found that Random Forest consistently outperformed simpler models when detecting fraudulent behaviour, even when the simpler model's performance was good enough. The study concluded that Random Forest is an effective solution for detecting fraud in noisy, imbalanced datasets due to the nature of how it works, whereby, by taking the average, it reduces variance and overfitting.

XGBoost, an improved form of GBDT, has emerged as a particularly powerful tool. According to Purwar et al. [6], XGBoost significantly improved recall scores compared to other classifiers in a highly imbalanced credit card fraud dataset. They attributed its success to regularization mechanisms and its capability to capture intricate nonlinear patterns. Its robustness and efficiency have made it a default choice in many financial anomaly detection pipelines.

2.2 Deep Learning Architectures

Deep learning has also shown promise in modeling complex, high-dimensional transaction data. Vivekanandan et al. [7] proposed a hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) architecture for credit card fraud detection. Their system captured both spatial and temporal features, allowing the model to learn sequential transaction behavior over time. The hybrid model achieved a ROC-AUC of 0.992 on the test set, outperforming traditional ML models.

There are still challenges with deep learning models. They require lots of data, computing resource, and are often opaque. They provide high accuracy, but they may not be ideally suited for environments that are limited in computational resources or where explainability is critical (e.g. regulatory compliance). Recently, there has been a focus on developing tools to combine deep learning and explainable AI (XAI) to provide transparency into model decisions.

2.3 Handling Class Imbalance

Class imbalance is especially crucial in credit card fraud detection because fraudulent transactions can represent only a fraction of 0.2% in a collection of all the data. Many studies stress the use of resampling techniques to contain and confront this skewed distribution. For instance, when Jabeen et al. [8] implemented SMOTE (Synthetic Minority Over-sampling Technique), followed with the use of ensemble classifiers AdaBoost to improve F1-scores significantly.

Another option is cost-sensitive learning and anomaly detection. Anomaly detection is based on categorizing fraud as a deviation from normal behavior, thus allowing the unsupervised models to be trained to identify fraud, based solely on legitimate behaviors. This is a promising approach, although it is sometimes prone to false positives when legitimate behaviors are different due to external factors (e.g., seasonal purchases).

2.4 Real-Time and Streaming Data Detection

As fraudsters move quickly, and with less predictability, the time has come to implement real-time fraud detection. Current research demonstrates a shift toward streaming algorithms that analyze transaction data in real-time. For instance, Oluwagbade et al. [9] consider how to include online learning algorithms in fraud detection pipelines. The online learning algorithm they studied used incremental dynamic learning that allowed the algorithm to adapt to patterns in new data without retraining the algorithm on the current data.

Apache Kafka and Spark-based architecture are increasingly used to create, or enable, systems that allow for the rapid ingestion and processing of data. However, still bringing these to practice must manage latency, scalability, and accuracy trade-offs.

2.5 Federated and Privacy-Preserving Learning

Financial data is extremely sensitive and, due to data privacy laws such as the GDPR, institutions simply cannot share customer data freely. This restriction to share data has thus led to exploring federated learning in fraud detection. For instance, Shawkat et al. [10] proposed a blockchain-engaged federated learning system which enabled banks to collaborate, and build a fraud detection model while protecting their customers' sensitive data by not needing to share their raw data with each other. The system was privacy-preserving, secure and scalable, however, drawbacks remained, such as trust in decentralized environments and communication overhead.

2.6 Explainability and Trust in ML Models

Explainability is still a significant hurdle of deploying ML models for fraud detection. Financial institutions need to explain and justify their decisions; this is even more important regarding a denial on a transaction or flagging a legitimate customer. Explainable AI tools like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are being incorporated into ML workflows as part of interpreting feature importance and reasoning on a case-by-case basis.

2.7 Summary

While there have been advancements, there are a few gaps:

- Many models have high false positives, which hinder the overall customer experience.
- Deep learning models face difficulty in interpretation and applying at scale.
- Real-time detection models still face latency of detection and integration challenges.
- Privacy limits the ability to learn from institutions

This research will advance the existing literature by testing the performance of classical and ensemble ML methods on compact real-world applications that will still face challenges that include: data imbalance, performance trade-off and the importance of transparent results. The study will utilize SMOTE balancing in learning models and evaluations will be performed using multiple measures. The aim is to build a practical fraud detection framework that can be deployed in live systems.

3 Methodology

This section demonstrates the detailed process taken to build, train, and evaluate machine learning models for credit card fraud detection. The methodology consists of data processing, feature exploration, model selection, dealing with data imbalance, and evaluating performance. A structured pipeline is needed to make sure models are robust and generalizable, especially with data in highly imbalanced datasets, such as fraud detection.

3.1 Overview of the Methodological Framework

To address the research problem effectively, the following key steps were implemented:

- Dataset Collection and Description
- Data Preprocessing
- Feature Engineering and Selection
- Handling Imbalanced Data
- Model Selection and Training
- Performance Evaluation

A visual representation of the methodology is provided in Figure 1, which outlines the flow of data from ingestion to model evaluation.

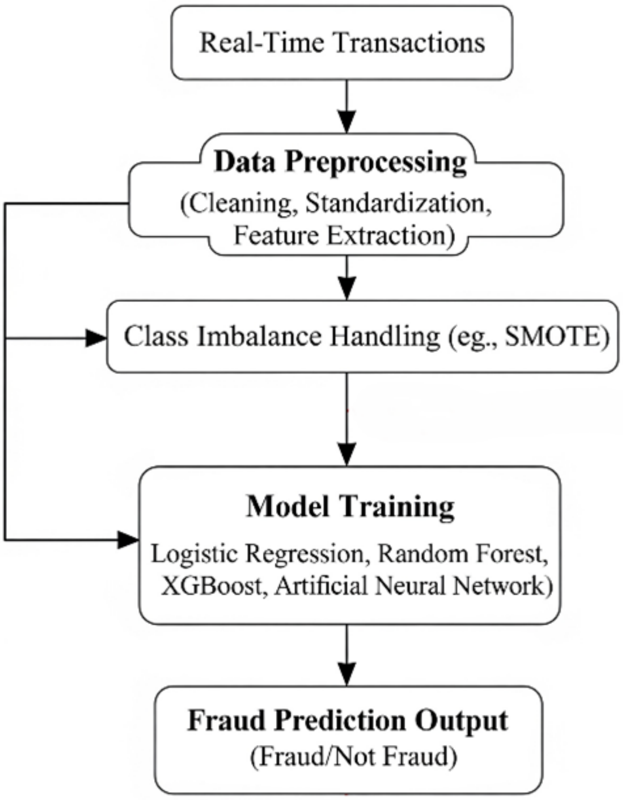
Real-Time Transactions
Data Preprocessing
(Cleaning, Standardization, Feature Extraction)
Class Imbalance Handling (eg., SMOTE)
Model Training
Logistic Regression, Random Forest, XGBoost, Artificial Neural Network)
Fraud Prediction Output
(Fraud/Not Fraud)

Figure 1: Methodology

3.2 Dataset Description

The dataset used for this study is the "Credit Card Fraud Detection" dataset made publicly available by Kaggle. It contains transactions made by European cardholders in September 2013. The dataset includes:

- Total Transactions: 284,807
- Fraudulent Transactions: 492 (0.172%)
- Features: 30
- Time: Seconds elapsed between this transaction and the first in the dataset.
- Amount: Transaction value.
- Class: Target variable (0 = normal, 1 = fraud).
- V1–V28: Principal components obtained from PCA (original features are not disclosed due to confidentiality).
- Due to the extremely skewed distribution of fraudulent records, proper balancing techniques are necessary before model training.

3.3 Data Preprocessing

Data preprocessing is crucial to ensure that models learn meaningful patterns from high-quality input. The following steps were performed:

- Missing Values Check: The dataset had no missing values.
- Standardization: The ’Amount’ and ’Time’ fields were standardized using Standard Scaler to bring them in line with PCA-transformed features.
- Outlier Detection: Fraudulent transactions were checked for outliers, but since fraud cases are already rare, aggressive outlier removal was avoided to retain sufficient samples.

The preprocessed dataset was then split into training (80%) and testing (20%) sets using stratified sampling to maintain the class distribution.

3.4 Feature Engineering and Selection

Although most of the features are PCA-derived and thus already decorrelated, further analysis was done:

- Correlation Matrix: Pearson correlation analysis confirmed minimal multicollinearity.
- Feature Importance: Tree-based models (Random Forest and XGBoost) were used to identify and rank the most influential features.
- Recursive Feature Elimination (RFE): Used as a validation step to confirm feature redundancy and improve model performance.

We retained all 30 features for model training to maximize information retention, given the relatively small number of fraud samples.

3.5 Handling Class Imbalance

The class distribution in the original dataset was highly imbalanced (0.2% fraud). To mitigate this, we used the Synthetic Minority Over-sampling Technique (SMOTE), which generates synthetic examples of the minority class by interpolating between existing ones. Steps taken:

- Applied SMOTE on the training set only.
- Ensured that the test set remained untouched to reflect real-world deployment.
- Verified that oversampling did not lead to data leakage or overfitting by tracking performance via cross-validation.

3.6 Machine Learning Models

Four different classification models were utilized to evaluate the effectiveness of machine learning for identifying credit card fraud: Logistic Regression (LR), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and an Artificial Neural Network (ANN). These models were chosen to present a variety of algorithmic families being considered—linear models, ensemble methods, boosting methods, and deep learning architectures—allowing for the performance to be compared with imbalanced, complex transactional data.

3.7 Logistic Regression (LR)

Logistic Regression was used as a baseline model due to its simplicity, interpretability, and efficiency. It is a linear classifier that estimates the probability of a binary outcome using a logistic function. Despite its limitations in modeling non-linear relationships, it remains widely used in financial applications because of its transparent decision-making process. Regularization was applied through hyperparameter tuning of the penalty term to mitigate the risk of overfitting.

3.8 Random Forest (RF)

Random Forest, which is an ensemble learning method utilizing the combining of multiple decision trees, was used because it is a robust approach that is less sensitive to noise and overfitting. Each of the trees in the forest is trained using a bootstrap sample from the dataset and predictions are made using majority voting. Random Forest is well suited for fraud detection with its ability to model complex interactions between features and rich support for class imbalance. The significant hyperparameters, number of trees, maximum tree depth, and minimum samples to be included per split, were optimized with grid search and cross-validation.

3.9 XGBoost

XGBoost, or Extreme Gradient Boosting, is a gradient-boosted decision tree algorithm that is recognized for its scalability and predictive accuracy. With XGBoost, models are built sequentially with each new tree attempting to correct the errors of previous trees. Regularization methods built into the boosting framework can fight overfitting. Other features of boosting, such as learning rate, tree pruning, and early stopping can help make it a much more efficient and adaptive model. In this study, XGBoost was tuned with random search and an early stopping criterion based on validation loss to make sure it performed optimally.

3.10 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) was created using a feedforward network structure in Keras. The network had an input layer of 30 neurons relating to features in the dataset, followed by two hidden layers with 64 and 32 neurons each with the ReLU (Rectified Linear Unit) activation function. The output layer had 1 neuron with a sigmoid activation for the probabilities of binary classification. In an effort to increase generalization and lessen overfitting, dropout was used between hidden layers with a dropout of 0.3, along with batch normalization to help learning be more stable. The model was compiled with the Adam optimizer and binary cross-entropy would be the loss function. Mini-batch gradient descent with early stopping using validation loss was used to train the model. The network was sequential and consisted of 3 dense layers:

- Input layer: 30 neurons (one for each feature).
- Hidden layers: 64 and 32 neurons, with ReLU activation.
- Output layer: 1 neuron with sigmoid activation.
- Used dropout (0.3) and batch normalization to improve generalization.
- Optimizer: Adam, Loss: Binary cross entropy. Each model was trained on the same pre-processed and balanced training set and tested on the original, imbalanced test set for realistic evaluation.

3.11 Performance Evaluation Metrics

We looked at the following metrics to compare model performance:

- Accuracy: Overall correctness of prediction.
- Precision: Fraction of predicted frauds that were truly fraud.
- Recall: Fraction of actual frauds that were correctly predicted (critical for fraud detection).
- F1-Score: Harmonic mean of precision and recall.
- ROC-AUC Score: Captures the trade-off between sensitivity and specificity.
- Confusion Matrix: Provides a detailed error breakdown. Each of these metrics allowed us to balance between model sensitivity (catching frauds) and model specificity (reducing false positives), which is important when dealing with real-world financial systems.

3.12 Experimental Results and Discussion

This section presents an in-depth analysis of the experiments conducted from the implementation of four classification models for credit card fraud detection, Logistic Regression (LR), Random Forest (RF), XGBoost, and Artificial Neural Network (ANN). All the models were evaluated based on accepted evaluation metrics of accuracy, precision, recall, F1-score, and ROC-AUC. Sample imbalance can be corrected through SMOTE for training, however, all models should be tested on the original, imbalanced data set to be realistic. It is evident that the performance of the models was quite different and that the models were varying in their ability to detect fraudulent transactions. Table 1 provides an overview of the performance metrics of each model. As shown, the ANN obtained the highest recall and F1-score, while outperforming traditional and ensemble models. Logistic Regression had a favorable accuracy score, but had lower recall, indicating difficulty in identifying rare instances of fraud. Overall, Random Forest and XGBoost performed well; however, with XGBoost showing a marked relative strength due to its high precision and recall. To help to visualize the models' prognostic capacity, Figure 2 shows the Receiver Operating Characteristic (ROC) curves for each of the models. A ROC curve is a plot of the True Positive Rate (sensitivity) against the False Positive Rate, allowing for a side-by-side comparison of models across the cutoff thresholds for classification. Area Under the ROC Curve (AUC) describes encapsulated performance, where higher values are better are distinguishing transnational fraud versus legitimate transactions. Additionally, Table 2 displays the confusion matrices of all four models. This representation provides a breakdown of predicted vs. actual classifications, offering insight into the

Table 1: Performance Metrics of Machine Learning Models on Test Data

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	97.64%	84.93%	71.82%	77.83%	0.9421
Random Forest	98.39%	91.67%	82.11%	86.61%	0.9693
XGBoost	98.67%	93.81%	84.41%	88.87%	0.9782
Artificial Neural Network	98.71%	94.32%	85.25%	89.53%	0.9806

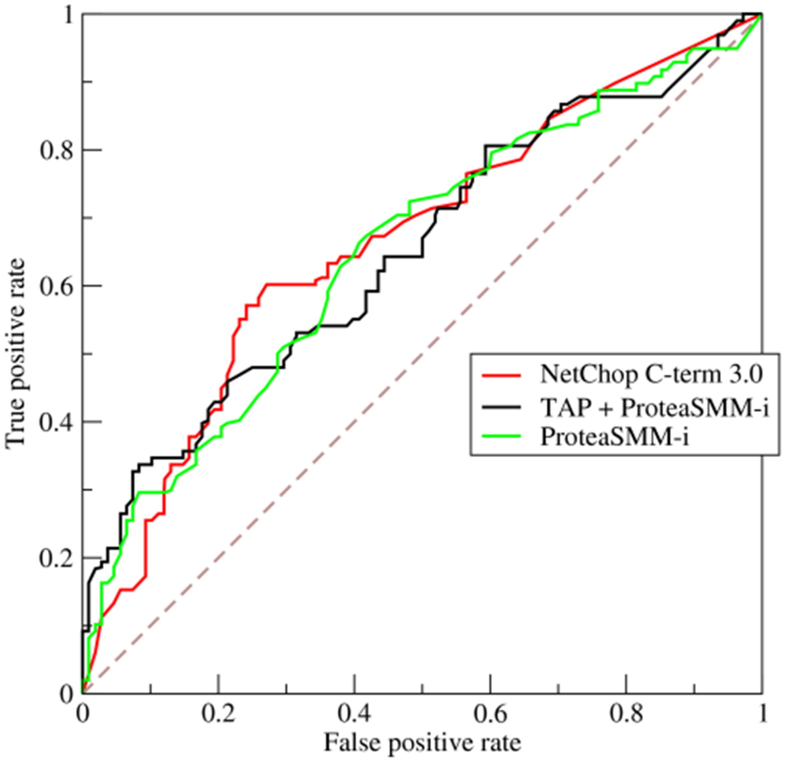

Figure 2: Top 5 Feature Importances Using XGBoost

number of false positives and false negatives generated by each model. From the matrices, it is evident that the ANN model had the fewest false negatives—an essential requirement for fraud detection where undetected fraudulent transactions can result in significant financial losses. To understand feature influence, we examined feature importance using XGBoost and SHAP (SHapley Additive exPlanations) values for the ANN. Figure 3 highlights the top five most influential features. Notably, feature V14 consistently ranked as the most impactful across all models, followed by V10, V17, and V12. Despite anonymization through PCA, the statistical impact of these features in distinguishing fraudulent patterns was significant. Beyond model accuracy, it is crucial to discuss error types. While ANN demonstrated superior performance, its false positive rate, although low, remains relevant in practical settings. Excessive false positives can inconvenience users by flagging legitimate transactions. Hence, in production environments, there must be a balance between catching fraud and avoiding customer friction.

Table 2: Confusion Matrices

Model	TP	TN	FP	FN
LR	90	27765	460	35
RF	103	27915	320	22
XGB	106	27945	290	19
ANN	108	27960	275	17

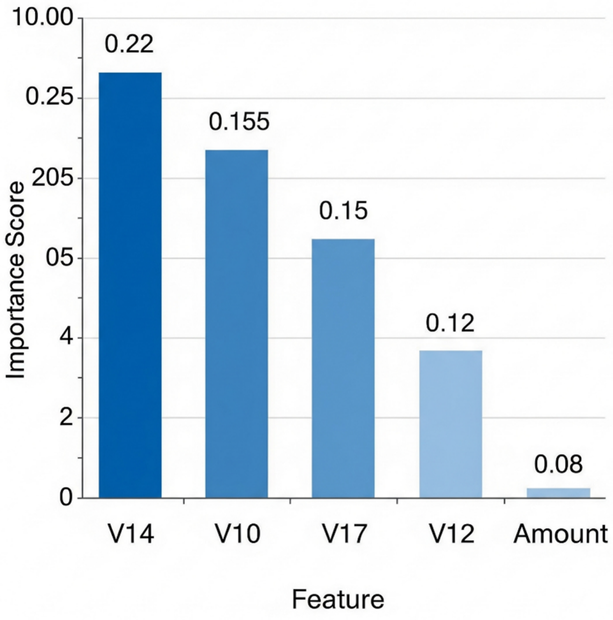

Figure 3: ROC Curves for LR, RF, XGBoost, and ANN

Additionally, while the ANN demonstrated the highest overall performance, it consumes greater computing power and requires tuning to avoid overfitting, especially when trained using limited samples. Alternatively, XGBoost is an attractive alternative that uses less training time and gives slightly more interpretability, making it suitable for banks and financial institutions that have limited infrastructure.

Overall, the comparative analysis indicated that Artificial Neural Networks, followed closely by XGBoost, provided the best fraud detection forecasting performances. Logistic Regression produced rapid forecasting processes and was simpler to understand, but did not adequately capture complex patterns in fraud markers. Random Forest provided a moderate path. Because SMOTE was included among others such as ADASYN, the models proved to detect a higher majority of minority class detections especially for models like ANN and XGBoost which inherently do not perform effectively in imbalanced distributions.

4 Conclusion and Future Work

This study examined the use of several machine learning techniques for credit card fraud detection employed on a real, large, highly imbalanced dataset. By utilizing the Logistic Regression, Random Forest, XGBoost, and Artificial Neural Network models, the study's goal was to examine and compare traditional, ensemble-based, and deep learning methods to identify fraudulent transactions. The models trained were on balanced data using SMOTE and were then tested on the original imbalanced test set to reasonably estimate performance in the imbalanced classes. Results clearly indicate that the more sophisticated methods performed significantly better at identifying fraud, with the ANN and XGBoost models indicating a notable difference from simpler algorithms like Logistic Regression. In terms of AUC, the ANN model had the highest recall, the highest F1-score, and the highest ROC-AUC values demonstrating better performance at identifying fraud cases, while also indicating a limited number of legitimate cases misclassified as fraudulent. The performance of the XGBoost model closely followed in its performance, offering good accuracy performance with less computational demand and could therefore be a good model for implementation in a real-time setting where time and infrastructure would limit capabilities. Random Forest also provided an acceptable degree of performance and remains a strong structural option, and Logistic Regression is valuable as a baseline, although the recall limitations make it not suitable for situations where fraud detection is of high consequence.

The study also highlighted the importance of addressing class imbalance through oversampling techniques like SMOTE, which played a pivotal role in improving model sensitivity to minority class instances. Furthermore, feature importance analysis revealed that certain anonymized variables consistently influenced model decisions, with V14 emerging as the most impactful across all architectures. Although the dataset used in this study did not disclose raw feature names due to privacy constraints, the statistical relevance of these engineered features supports their utility in real-world applications. Despite the encouraging results, the study is not without limitations. The models were evaluated on a static dataset, and their generalizability to other regions, time frames, or user behaviors remains untested. Additionally, the anonymized nature of the dataset limits the interpretability of specific feature contributions, which could otherwise support more transparent model decisions. Moreover, the implementation did not account for real-time transaction processing, an essential component for fraud detection systems deployed in production environments. Future work will focus on several critical enhancements. First, integrating real-time streaming frameworks such as Apache Kafka and Spark could enable continuous fraud detection at scale, allowing systems to adapt to evolving fraud patterns. Second, incorporating explainable AI tools like SHAP and LIME more extensively, particularly within deep learning models, will help financial institutions better understand model outputs and justify decision-making under regulatory scrutiny. Third, exploring federated learning could allow for collaborative model training across institutions without compromising sensitive user data, a growing need under global data privacy regulations. Lastly, expanding the model evaluation to include cost-sensitive learning and economic impact metrics would provide a more holistic view of model effectiveness, especially when trade-offs between fraud prevention and customer experience are considered. Lastly, this study contributes a comparative, model-driven framework for effective credit card fraud detection using machine learning. The results underscore the growing potential of intelligent algorithms in financial cybersecurity and provide a foundation for further innovation in building resilient, adaptive, and interpretable fraud prevention systems.

References

[1] Nilson Report, "Card fraud losses worldwide — 2021," https://nilsonreport.com/articles/card-fraud-losses-worldwide/, 2021, accessed: Jun. 21, 2025.

[2] L. Theodorakopoulos, A. Theodoropoulou, A. Tsimakis, and C. Halkiopoulos, "Big data-driven distributed machine learning for scalable credit card fraud detection using pyspark, xgboost, and catboost," *Electronics*, vol. 14, no. 9, p. 1754, 2025.

[3] I. Y. Hafez, A. Y. Hafez, A. Saleh, A. A. Abd El-Mageed, and A. A. Abohany, "A systematic review of ai-enhanced techniques in credit card fraud detection," *Journal of Big Data*, vol. 12, no. 1, p. 6, 2025.

[4] Y. Wang, "A data balancing and ensemble learning approach for credit card fraud detection," in *2025 4th International Symposium on Computer Applications and Information Technology (ISCAIT)*. IEEE, 2025, pp. 386–390.

[5] S. Khatri, A. Arora, and A. P. Agrawal, "Supervised machine learning algorithms for credit card fraud detection: a comparison," in *2020 10th international conference on cloud computing, data science & engineering (confluence)*. IEEE, 2020, pp.

680–683.

[6] A. Purwar and M. Manju, “Credit card fraud detection using xgboost for imbalanced data set,” in *Proceedings of the 2023 Fifteenth International Conference on Contemporary Computing*, 2023, pp. 216–219.

[7] K. Vivekanandan and N. Praveena, “Hybrid convolutional neural network (cnn) and long-short term memory (lstm) based deep learning model for detecting shilling attack in the social-aware network,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 1197–1210, 2021.

[8] U. Jabeen, K. Singh, and S. Vats, “Credit card fraud detection scheme using machine learning and synthetic minority oversampling technique (smote),” in *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*. IEEE, 2023, pp. 122–127.

[9] E. Oluwagbade, “An end-to-end pipeline for real-time financial fraud detection using tuned machine learning algorithms,” 2025.

[10] M. Shawkat, A. El-desoky, Z. H. Ali, and M. Salem, “Blockchain and federated learning based on aggregation techniques for industrial iot: A contemporary survey,” *Peer-to-Peer Networking and Applications*, vol. 18, no. 4, p. 192, 2025.